



网络存储作业 2

RAID

罗暄澍

学号：2015211527

班级：2015211314



一、有了 RAID 我们还需要备份系统吗？

我认为依然需要。这个结论基于以下几个原因：

1、RAID 是一种磁盘冗余机制，而不是备份

RAID 最初的研制目的是为了组合小的廉价磁盘来代替大的昂贵磁盘，以降低大批量数据存储的费用，同时也希望采用冗余信息的方式，使得磁盘失效时不会使对数据的访问受损失，从而开发出一定水平的数据保护技术。磁盘阵列是由很多价格较便宜的磁盘所组合成的一个容量巨大的磁盘组，利用个别磁盘提供数据所产生加成效果提升整个磁盘系统效能。利用这项技术，将数据切割成许多区段，分别存放在各个硬盘上。因此，RAID 严格意义上来说是一种存储方式。

也就是说，RAID 完成的是一个“1+1>2”的过程，相对于多个独立的硬盘存储，通过 RAID 技术，在损失一定磁盘利用率的情况下，得到了更好的读写速度和安全性。RAID 本身只是一种磁盘冗余机制，是一种存储方式。相对于读写性能的绝对提高，RAID 对数据安全的保障只是相对的。

2、RAID 技术承诺的安全性是有限的

不同的 RAID 技术有不同的安全等级。比如说 RAID0 本身并不能保障数据安全，RAID1 由于有备份盘的存在，可以承受一块硬盘损坏或者不同组的多块硬盘同时损坏，但是如果同组的两块硬盘同时损坏（数据盘和对应的镜像盘），数据依然无法恢复，其他 RAID 技术基本同理。也就是说 RAID 的数据安全是有限水平的、相对的，不是绝对稳健的。而如果有整体备份的系统，就算所有硬盘全坏了，数据依然可以恢复。单纯依靠 RAID 的系统安全性和带有备份的系统安全性是不可同日而语的。

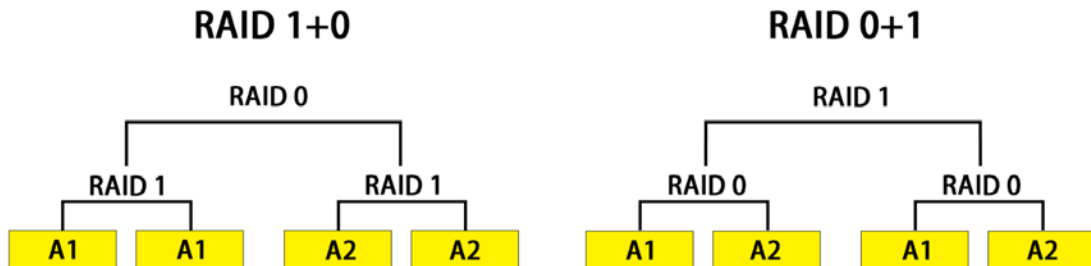
3、RAID 的安全性是建立在 RAID 系统能够正常运转的前提下的

RAID 体系在正常运转的时候，可以提供一定程度的安全性。但是由于管理员操作失误、黑客入侵或者病毒感染，整个体系可能崩溃，所有硬盘的数据都会消失，到时仅靠 RAID 体系则毫无用处。

综上，RAID 是一种磁盘冗余机制，而非备份机制，RAID 只是具备一定水平的保障数据安全的能力，无法面对灾难性的事件。所以即使使用了 RAID 技术，系统的完整备份仍然十分有必要。

二、说明 RAID0+1 和 RAID1+0 谁更好(以 4 块具有相同可靠度硬盘为例说明)

具有四块硬盘的 RAID 0+1 和 RAID 1+0 体系如下图所示：



按照 PPT 中的评价标准进行比较

1、最少磁盘数

显而易见都是四块硬盘

2、磁盘利用率

由于镜像盘的存在均为 50%

3、开销

两种策略的开销相同

4、读性能（较单磁盘）

相同，均快于单磁盘

5、写性能（较单磁盘）

RAID 0+1 优于 RAID 1+0。因为 RAID 0+1 先进行分段，RAID 1+0 则需要先进行备份，若一段数据写入一块单磁盘的时间为 t ，那么 RAID 0+1 仅需 $t/2$ 的时间，RAID 1+0 体系则会略慢于 t （毕竟还有数据复制的时间）。

6、数据恢复时间

RAID 1+0 优于 RAID 0+1。在 RAID 1+0 体系下，损坏的硬盘可以快速地通过底层的 RAID 1 进行恢复。而当 RAID 0+1 体系中出现硬盘损坏的情况，级数则下降成 RAID 0，数据恢复相对慢。

注：只有四块硬盘的话，容错率也是相同的，不做讨论。

可靠性计算比较

设每个硬盘的可靠性为 0.9

RAID 0+1 : $R = 1 - (1 - 0.9^2)^2 = 0.9639$

RAID 1+0 : $R = [1 - (1 - 0.9)(1 - 0.9)]^2 = 0.9801$

所以 RAID 1+0 更可靠。

结论

将两种体系不同点列出：

	写性能	数据恢复时间	可靠性
RAID 0+1	强	弱	弱
RAID 1+0	弱	强	强

三、在什么情况下 RAID1 + 0 比 RAID5 更适用？

因为 RAID 1+0 先组合 RAID 1 阵列，然后组成 RAID 0 模式，所以它有更好的容错能力。而 RAID5 是针对奇偶校验驱动器所带来的瓶颈而产生的解决方案，RAID 5 的写速度相对更慢。故而：

- 1、小 I/O 的数据库类型操作，RAID 1+0 在此时比 RAID 5 更为实用。所以典型的 OLTP 环境里，RAID 1+0 更好些。因为 OLTP 环境中，IO 性能上我们考虑的主要方面。
- 2、安全性要求高，不计成本的情况下，RAID 1+0 同样更为适用。

四、我们还可以如何提高 RAID 的性能？

1、根据需求

比如当对安全性要求不高，对速度要求高的时候，就要使用 RAID0，而无需校验和冗余，从而提升性能。反之，如果对安全性要求高，那么冗余多，有校验的策略就要使用，当然这样提升的是安全性能，而非速度方面的性能。

另一方面，是随机读写多还是连续读写多，也是一个影响因素。这涉及到 RAID 3 和 RAID 4、RAID 5 和 RAID 6 的选取。

这一点主要是对于相同硬件的处理方式角度说的，提升的是相对性能而非绝对性能。

2、RAID 卡和硬盘的选取

若要提升 RAID 的绝对性能，就要提高部件的绝对性能。比如 RAID 卡要选取硬 RAID 卡，而非软件实现，选取 SCSI 或 SAS RAID 卡，硬盘选取转速高、缓存高、读写速度快的。或者采用混合 RAID。总之在预算范围内，选取性能匹配的、更优秀的硬件，是提高性能的保障。

3、是否支持高级模式

高级模式支持指定 SSD 作为 RAID 的缓存，当 RAID 上有数据库时，特点数据的缓存放在 SSD 上，效果非常明显